

OIIPE

#2

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/922,501

DATE: 09/13/2001

TIME: 11:27:57

Input Set : A:\SR4877.txt

Output Set: N:\CRF3\08132001\I922501.raw

ENTERED

4 <110> APPLICANT: Dartois, Veronique A.  
 5 Hoch, James A.  
 6 Valle, Fernando  
 7 Kumar, Manoj

11 <120> TITLE OF INVENTION: 2, 5-DKG PERMEASES  
 14 <130> FILE REFERENCE: P-SR 4877

C--> 16 <140> CURRENT APPLICATION NUMBER: US/09/922,501  
 C--> 16 <141> CURRENT FILING DATE: 2001-08-03  
 16 <150> PRIOR APPLICATION NUMBER: US 09/633,294  
 17 <151> PRIOR FILING DATE: 2000-08-04  
 19 <150> PRIOR APPLICATION NUMBER: US 09/677,032  
 20 <151> PRIOR FILING DATE: 2000-09-29  
 22 <160> NUMBER OF SEQ ID NOS: 22  
 24 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 26 <210> SEQ ID NO: 1  
 27 <211> LENGTH: 1500  
 28 <212> TYPE: DNA  
 29 <213> ORGANISM: Unknown  
 31 <220> FEATURE:  
 32 <223> OTHER INFORMATION: environmental source  
 34 <221> NAME/KEY: CDS  
 35 <222> LOCATION: (94)...(1374)  
 37 <400> SEQUENCE: 1

38 ggcgaatagc cgggccggcg tcataataac ggccttctct gtaccctaca tacggcgcg 60  
 39 gcgtcatgaa cctcaacttt agtaggcaag cct atg aac agc tct acc aat gca 114  
 40 Met Asn Ser Ser Thr Asn Ala  
 41 1 5  
 43 acg aaa cgc tgg tgg tac atc atg cct atc gtg ttt atc acg tat agc 162  
 44 Thr Lys Arg Trp Trp Tyr Ile Met Pro Ile Val Phe Ile Thr Tyr Ser  
 45 10 15 20  
 47 ctg gcg tat ctc gac cgc gca aac ttc agc ttt gct tcg gca gcg ggc 210  
 48 Leu Ala Tyr Leu Asp Arg Ala Asn Phe Ser Phe Ala Ser Ala Ala Gly  
 49 25 30 35  
 51 att acg gaa gat tta ggc att acc aaa ggc atc tcg tcg ctt ctt ggc 258  
 52 Ile Thr Glu Asp Leu Gly Ile Thr Lys Gly Ile Ser Ser Leu Leu Gly  
 53 40 45 50 55  
 55 gca ctt ttc ttc ctc ggc tat ttc ttc ttc cag atc ccg ggg gcg att 306  
 56 Ala Leu Phe Phe Leu Gly Tyr Phe Phe Phe Gln Ile Pro Gly Ala Ile  
 57 60 65 70  
 59 tac gcg gaa cgc cgt agc gta cgg aag ctg att ttc atc tgt ctg atc 354  
 60 Tyr Ala Glu Arg Arg Ser Val Arg Lys Leu Ile Phe Ile Cys Leu Ile  
 61 75 80 85  
 63 ctg tgg ggc gcc tgc gcc tcg ctt gac cgg gat ggt gca caa tat tcc 402  
 64 Leu Trp Gly Ala Cys Ala Ser Leu Asp Arg Asp Gly Ala Gln Tyr Ser  
 65 90 95 100  
 67 agc gct ggc tgg cga tcc gtt tta ttc tcg gct gtc gtg gaa gcg gcg 450  
 68 Ser Ala Gly Trp Arg Ser Val Leu Phe Ser Ala Val Val Glu Ala Ala

## RAW SEQUENCE LISTING

DATE: 08/13/2001

PATENT APPLICATION: US/09/922,501

TIME: 11:27:57

Input Set : A:\SR4877.txt

Output Set: N:\CRF3\08132001\I922501.raw

69	105	110	115	
71	gtc atg ccg gcg atg ctg att tac atc agt aac tgg ttt acc aaa tca	498		
72	Val Met Pro Ala Met Leu Ile Tyr Ile Ser Asn Trp Phe Thr Lys Ser			
73	120 125 130 135			
75	gaa cgt tca cgc gcc aac acc ttc tta atc ctc ggc aac ccg gtc acg	546		
76	Glu Arg Ser Arg Ala Asn Thr Phe Leu Ile Leu Gly Asn Pro Val Thr			
77	140 145 150			
79	gta ctg tgg atg tcg gtg gtc tcc ggc tac ctg att cag tcc ttc ggc	594		
80	Val Leu Trp Met Ser Val Val Ser Gly Tyr Leu Ile Gln Ser Phe Gly			
81	155 160 165			
83	tgg cgt gaa atg ttt att att gaa ggc gtt ccg gcc gtc ctc tgg gcc	642		
84	Trp Arg Glu Met Phe Ile Ile Glu Gly Val Pro Ala Val Leu Trp Ala			
85	170 175 180			
87	ttc tgc tgg tgg gtg ctg gtc aaa gtt aaa ccg tcg cag gtg aac tgg	690		
88	Phe Cys Trp Trp Val Leu Val Lys Val Lys Pro Ser Gln Val Asn Trp			
89	185 190 195			
91	ttg tcg gaa aac gag aaa gcc gcg ctg cag gcg cag ctg gag agc gag	738		
92	Leu Ser Glu Asn Glu Lys Ala Ala Leu Gln Ala Gln Leu Glu Ser Glu			
93	200 205 210 215			
95	cag cag ggt att aaa gcc gtg cgt aac tac ggc gaa gcc ttc cgc tca	786		
96	Gln Gln Gly Ile Lys Ala Val Arg Asn Tyr Gly Glu Ala Phe Arg Ser			
97	220 225 230			
99	cgc aac gtc att cta ctg tgc atg cag tat ttt gcc tgg agt atc ggc	834		
100	Arg Asn Val Ile Leu Leu Cys Met Gln Tyr Phe Ala Trp Ser Ile Gly			
101	235 240 245			
103	gtg tac ggt ttt gtg ctg tgg ttg ccg tca att att cgc agc ggc ggc	882		
104	Val Tyr Gly Phe Val Leu Trp Leu Pro Ser Ile Ile Arg Ser Gly Gly			
105	250 255 260			
107	gtc aat atg ggg atg gtg gaa gtc ggc tgg ctc tct tcg gtg cct tat	930		
108	Val Asn Met Gly Met Val Glu Val Gly Trp Leu Ser Ser Val Pro Tyr			
109	265 270 275			
111	ctg gcc gcg act att gcg atg atc gtc gtc tcc tgg gct tcc gat aaa	978		
112	Leu Ala Ala Thr Ile Ala Met Ile Val Val Ser Trp Ala Ser Asp Lys			
113	280 285 290 295			
115	atg cag aac cgt aaa ctg ttc gtc tgg ccg ctg ctg ctg att ggc gga	1026		
116	Met Gln Asn Arg Lys Leu Phe Val Trp Pro Leu Leu Leu Ile Gly Gly			
117	300 305 310			
119	ctg gct ttt att ggc tca tgg gcc gtc ggc gct aac cat ttc tgg gcc	1074		
120	Leu Ala Phe Ile Gly Ser Trp Ala Val Gly Ala Asn His Phe Trp Ala			
121	315 320 325			
123	tct tat acc ctg ctg gtg att gcc aat gcg gca atg tac gcc cct tac	1122		
124	Ser Tyr Thr Leu Leu Val Ile Ala Asn Ala Ala Met Tyr Ala Pro Tyr			
125	330 335 340			
127	ggt ccg ttt ttc gcc atc att ccg gaa atg ctg ccg cgt aac gtc gcc	1170		
128	Gly Pro Phe Phe Ala Ile Ile Pro Glu Met Leu Pro Arg Asn Val Ala			
129	345 350 355			
131	ggt ggc gca atg gcg ctc atc aac agc atg ggg gcc tta ggt tca ttc	1218		
132	Gly Gly Ala Met Ala Leu Ile Asn Ser Met Gly Ala Leu Gly Ser Phe			
133	360 365 370 375			

## RAW SEQUENCE LISTING

DATE: 08/13/2001

PATENT APPLICATION: US/09/922,501

TIME: 11:27:57

Input Set : A:\SR4877.txt

Output Set: N:\CRF3\08132001\I922501.raw

```

135 ttt ggt tcg tgg ttc gtg ggc tac ctg aac ggc acc acc ggc agt cca 1266
136 Phe Gly Ser Trp Phe Val Gly Tyr Leu Asn Gly Thr Thr Gly Ser Pro
137          380          385          390
139 tca gcc tca tac att ttc atg gga gtg gcg ctt ttc gcc tcg gta tgg 1314
140 Ser Ala Ser Tyr Ile Phe Met Gly Val Ala Leu Phe Ala Ser Val Trp
141          395          400          405
143 ctt act tta att gtt aag cct gct aac aat caa aag ctc ccc atc ggc 1362
144 Leu Thr Leu Ile Val Lys Pro Ala Asn Asn Gln Lys Leu Pro Ile Gly
145          410          415          420
147 gct cgt cac gcc tgacctttac tacttacgga gatcacgcct tgggtacgtt 1414
148 Ala Arg His Ala
149          425
151 gcaggacaaa ccgataggca ccgcaaaggc tggggccatc gagcagcgcg taaacagtca 1474
152 gctggttgct gtcgctgtgc ggcgtc 1500
154 <210> SEQ ID NO: 2
155 <211> LENGTH: 427
156 <212> TYPE: PRT
157 <213> ORGANISM: Unknown
159 <220> FEATURE:
160 <223> OTHER INFORMATION: environmental source
162 <400> SEQUENCE: 2
163 Met Asn Ser Ser Thr Asn Ala Thr Lys Arg Trp Trp Tyr Ile Met Pro
164 1          5          10          15
165 Ile Val Phe Ile Thr Tyr Ser Leu Ala Tyr Leu Asp Arg Ala Asn Phe
166          20          25          30
167 Ser Phe Ala Ser Ala Ala Gly Ile Thr Glu Asp Leu Gly Ile Thr Lys
168          35          40          45
169 Gly Ile Ser Ser Leu Leu Gly Ala Leu Phe Phe Leu Gly Tyr Phe Phe
170          50          55          60
171 Phe Gln Ile Pro Gly Ala Ile Tyr Ala Glu Arg Arg Ser Val Arg Lys
172 65          70          75          80
173 Leu Ile Phe Ile Cys Leu Ile Leu Trp Gly Ala Cys Ala Ser Leu Asp
174          85          90          95
175 Arg Asp Gly Ala Gln Tyr Ser Ser Ala Gly Trp Arg Ser Val Leu Phe
176          100          105          110
177 Ser Ala Val Val Glu Ala Ala Val Met Pro Ala Met Leu Ile Tyr Ile
178          115          120          125
179 Ser Asn Trp Phe Thr Lys Ser Glu Arg Ser Arg Ala Asn Thr Phe Leu
180          130          135          140
181 Ile Leu Gly Asn Pro Val Thr Val Leu Trp Met Ser Val Val Ser Gly
182 145          150          155          160
183 Tyr Leu Ile Gln Ser Phe Gly Trp Arg Glu Met Phe Ile Ile Glu Gly
184          165          170          175
185 Val Pro Ala Val Leu Trp Ala Phe Cys Trp Trp Val Leu Val Lys Val
186          180          185          190
187 Lys Pro Ser Gln Val Asn Trp Leu Ser Glu Asn Glu Lys Ala Ala Leu
188          195          200          205
189 Gln Ala Gln Leu Glu Ser Glu Gln Gln Gly Ile Lys Ala Val Arg Asn
190          210          215          220

```

## RAW SEQUENCE LISTING

DATE: 08/13/2001

PATENT APPLICATION: US/09/922,501

TIME: 11:27:57

Input Set : A:\SR4877.txt

Output Set: N:\CRF3\08132001\I922501.raw

```

191 Tyr Gly Glu Ala Phe Arg Ser Arg Asn Val Ile Leu Leu Cys Met Gln
192 225                230                235                240
193 Tyr Phe Ala Trp Ser Ile Gly Val Tyr Gly Phe Val Leu Trp Leu Pro
194                245                250                255
195 Ser Ile Ile Arg Ser Gly Gly Val Asn Met Gly Met Val Glu Val Gly
196                260                265                270
197 Trp Leu Ser Ser Val Pro Tyr Leu Ala Ala Thr Ile Ala Met Ile Val
198                275                280                285
199 Val Ser Trp Ala Ser Asp Lys Met Gln Asn Arg Lys Leu Phe Val Trp
200                290                295                300
201 Pro Leu Leu Leu Ile Gly Gly Leu Ala Phe Ile Gly Ser Trp Ala Val
202 305                310                315                320
203 Gly Ala Asn His Phe Trp Ala Ser Tyr Thr Leu Leu Val Ile Ala Asn
204                325                330                335
205 Ala Ala Met Tyr Ala Pro Tyr Gly Pro Phe Phe Ala Ile Ile Pro Glu
206                340                345                350
207 Met Leu Pro Arg Asn Val Ala Gly Gly Ala Met Ala Leu Ile Asn Ser
208                355                360                365
209 Met Gly Ala Leu Gly Ser Phe Phe Gly Ser Trp Phe Val Gly Tyr Leu
210                370                375                380
211 Asn Gly Thr Thr Gly Ser Pro Ser Ala Ser Tyr Ile Phe Met Gly Val
212 385                390                395                400
213 Ala Leu Phe Ala Ser Val Trp Leu Thr Leu Ile Val Lys Pro Ala Asn
214                405                410                415
215 Asn Gln Lys Leu Pro Ile Gly Ala Arg His Ala
216                420                425
219 <210> SEQ ID NO: 3
220 <211> LENGTH: 1775
221 <212> TYPE: DNA
222 <213> ORGANISM: Unknown
224 <220> FEATURE:
225 <223> OTHER INFORMATION: environmental source
227 <221> NAME/KEY: CDS
228 <222> LOCATION: (214)...(1491)
230 <400> SEQUENCE: 3
231 ggcaatttgc ggtgtttttt cgcaggacg ttcacgtgcc ggctgtatt catcaacggc 60
232 cctgcgctat tcgcaaagtg gtggtgaaaa taccgctgcg ttatttaacg cccaataagc 120
233 aacaccgagt ttataaccct gaacgacacg gctgcgggcc tgtgtagacg cccctacgcc 180
234 ttaacaccac taaatgactc tacaggtgta tat atg aat aca gcc tct gtt tct 234
235                Met Asn Thr Ala Ser Val Ser
236                1                5
238 gtc acc caa agc cag gcg atc ccc aaa tta cgc tgg ttg aga ata gtg 282
239 Val Thr Gln Ser Gln Ala Ile Pro Lys Leu Arg Trp Leu Arg Ile Val
240                10                15                20
242 ccg cct att ctt att acc tgc att att tcc tat atg gac cgg gtg aac 330
243 Pro Pro Ile Leu Ile Thr Cys Ile Ile Ser Tyr Met Asp Arg Val Asn
244                25                30                35
246 atc gcc ttc gcc atg ccc ggc ggc atg gac gat gaa ctg ggc atc acc 378
247 Ile Ala Phe Ala Met Pro Gly Gly Met Asp Asp Glu Leu Gly Ile Thr

```

## RAW SEQUENCE LISTING

DATE: 08/13/2001

PATENT APPLICATION: US/09/922,501

TIME: 11:27:57

Input Set : A:\SR4877.txt

Output Set: N:\CRF3\08132001\I922501.raw

248	40		45		50		55	
250	gcc	tcg	atg	gcc	ggg	ttg	gcc	ggc
251	Ala	Ser	Met	Ala	Gly	Leu	Ala	Gly
252								
254	ttc	ttg	cag	gta	ccc	ggc	ggc	aag
255	Phe	Leu	Gln	Val	Pro	Gly	Gly	Lys
256								
258	aaa	ttc	atc	ggt	tgg	tcg	ttg	ttg
259	Lys	Phe	Ile	Gly	Trp	Ser	Leu	Leu
260								
262	acc	ggg	ctg	gtc	acg	aat	cag	tat
263	Thr	Gly	Leu	Val	Thr	Asn	Gln	Tyr
264								
266	ctc	ggc	cgt	ttc	cga	agc	ggc	atg
267	Leu	Gly	Arg	Phe	Arg	Ser	Gly	Met
268	120							
270	agc	aac	tgg	ttc	ccg	gac	aag	gaa
271	Ser	Asn	Trp	Phe	Pro	Asp	Lys	Glu
272								
274	atc	atg	ttc	gtg	ccg	atc	gcc	ggc
275	Ile	Met	Phe	Val	Pro	Ile	Ala	Gly
276								
278	tgg	atc	atc	acc	gcc	tgg	gac	tgg
279	Trp	Ile	Ile	Thr	Ala	Trp	Asp	Trp
280								
282	gcg	ctg	tcg	ctg	gtc	gtg	atg	gtg
283	Ala	Leu	Ser	Leu	Val	Val	Met	Val
284								
286	cgt	cca	caa	gag	gcc	aaa	agg	att
287	Arg	Pro	Gln	Glu	Ala	Lys	Arg	Ile
288	200							
290	atc	aaa	acg	ctg	cac	gac	gaa	cag
291	Ile	Lys	Thr	Leu	His	Asp	Glu	Gln
292								
294	cgc	aac	gcc	tcg	ctg	cgt	cgg	gtg
295	Arg	Asn	Ala	Ser	Leu	Arg	Arg	Val
296								
298	ttc	ttc	tac	cag	acc	ggg	ata	tac
299	Phe	Phe	Tyr	Gln	Thr	Gly	Ile	Tyr
300								
302	att	ctc	aag	ggg	ctc	acc	aac	ggc
303	Ile	Leu	Lys	Gly	Leu	Thr	Asn	Gly
304								
306	gct	atc	ctg	ccc	tat	atc	ggc	gcc
307	Ala	Ile	Leu	Pro	Tyr	Ile	Gly	Ala
308	280							
310	acc	ctc	tcc	gat	cgc	acc	ggc	aag
311	Thr	Leu	Ser	Asp	Arg	Thr	Gly	Lys
312								

VERIFICATION SUMMARY

DATE: 08/13/2001

PATENT APPLICATION: US/09/922,501

TIME: 11:27:58

Input Set : A:\SR4877.txt

Output Set: N:\CRF3\08132001\I922501.raw

L:16 M:270 C: Current Application Number differs, Replaced Current Application No

L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date